RSAC 2025 Conference

Many Voices. One Community.

SESSION ID: LAB3-T01

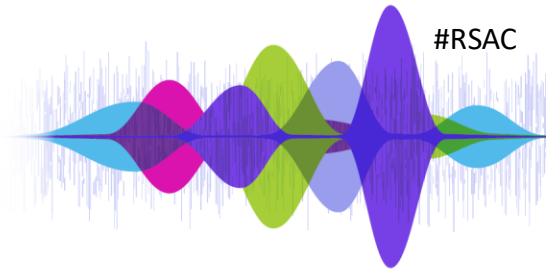
Offensive AI: Penetration Testing with Python and Artificial Intelligence

Tennisha Martin

Executive Director and Founder
BlackGirlsHack
https://www.linkedin.com/in/tennisha/

Aleise McGowan

CISO
BlackGirlsHack
https://www.linkedin.com/in/aleisemcgowan/



Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the presenters individually and, unless expressly stated to the contrary, are not the opinion or position of RSA Conference LLC or any other co-sponsors. RSA Conference LLC does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented.

Attendees should note that sessions may be audio- or video-recorded and may be published in various media, including print, audio and video formats without further notice. The presentation template and any media capture are subject to copyright protection.

© 2025 RSA Conference LLC or its affiliates. The RSAC and RSAC CONFERENCE logos and other trademarks are proprietary. All rights reserved.

Who Am I

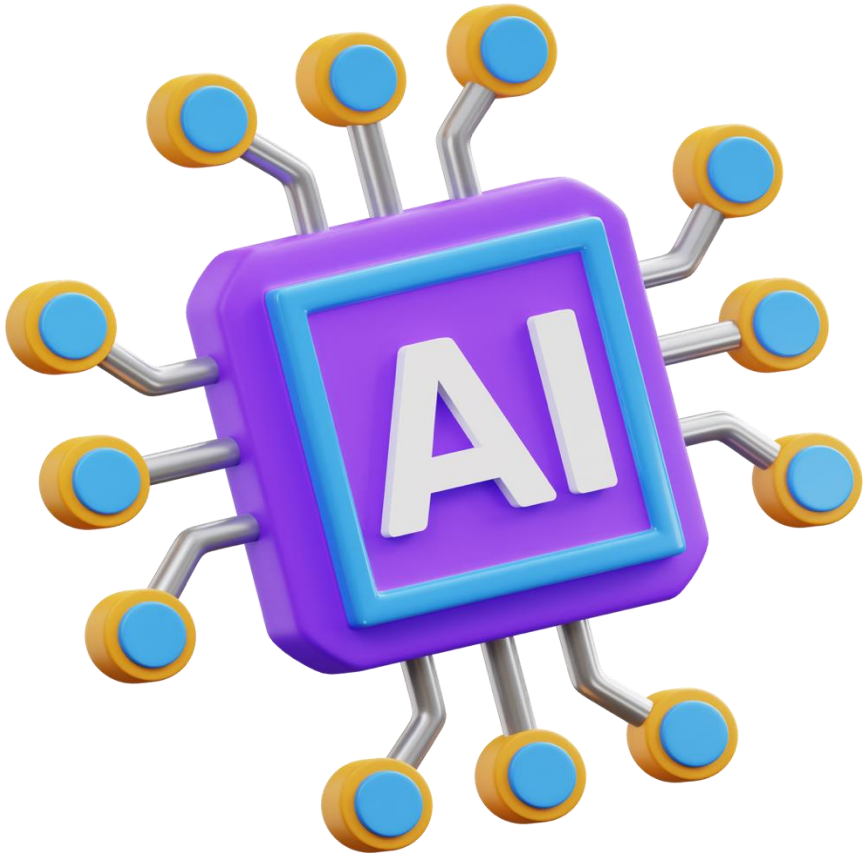
- Tennisha Martin
  - Executive Director/Founder BlackGirlsHack
  - Wife
  - Hacker
  - MSx4, MBA, GWAPT, GPEN, GCIH, Doctorate Loading
  - CEO BGH Security
  - Researcher of AI/Cyber Education/Offensive Security

Who Am I

- Aleise McGowan
  - CISO BlackGirlsHack
  - Wife
  - Hacker
  - PhD, CISSP, GCIH, GSNA, CYSA+, (GICSP and GWAPT Loading)
  - Assistant Professor University of Southern Mississippi
  - Researcher of AI/Cyber Education/Offensive Security/IT OT Security

Contents

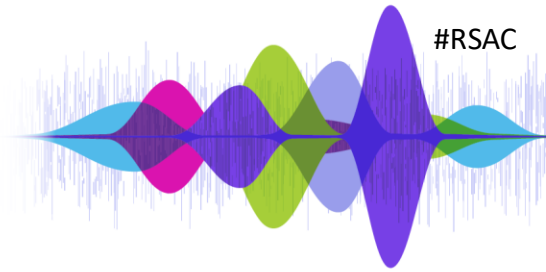
1. What is AI
2. What is Offensive AI?
3. What is Red Teaming?
4. Types of Artificial Intelligence
5. Current State
6. Future of Offensive AI



What is Offensive security?

- Red Teaming is a structured testing process that involves ethical hackers acting as a malicious actor and attempting emulate tactics, techniques and procedures (TTPs) against a company's systems.
- Red teams leverage attack methodologies to assess vulnerabilities, identify technical flaws, and provide recommendations to improve security.

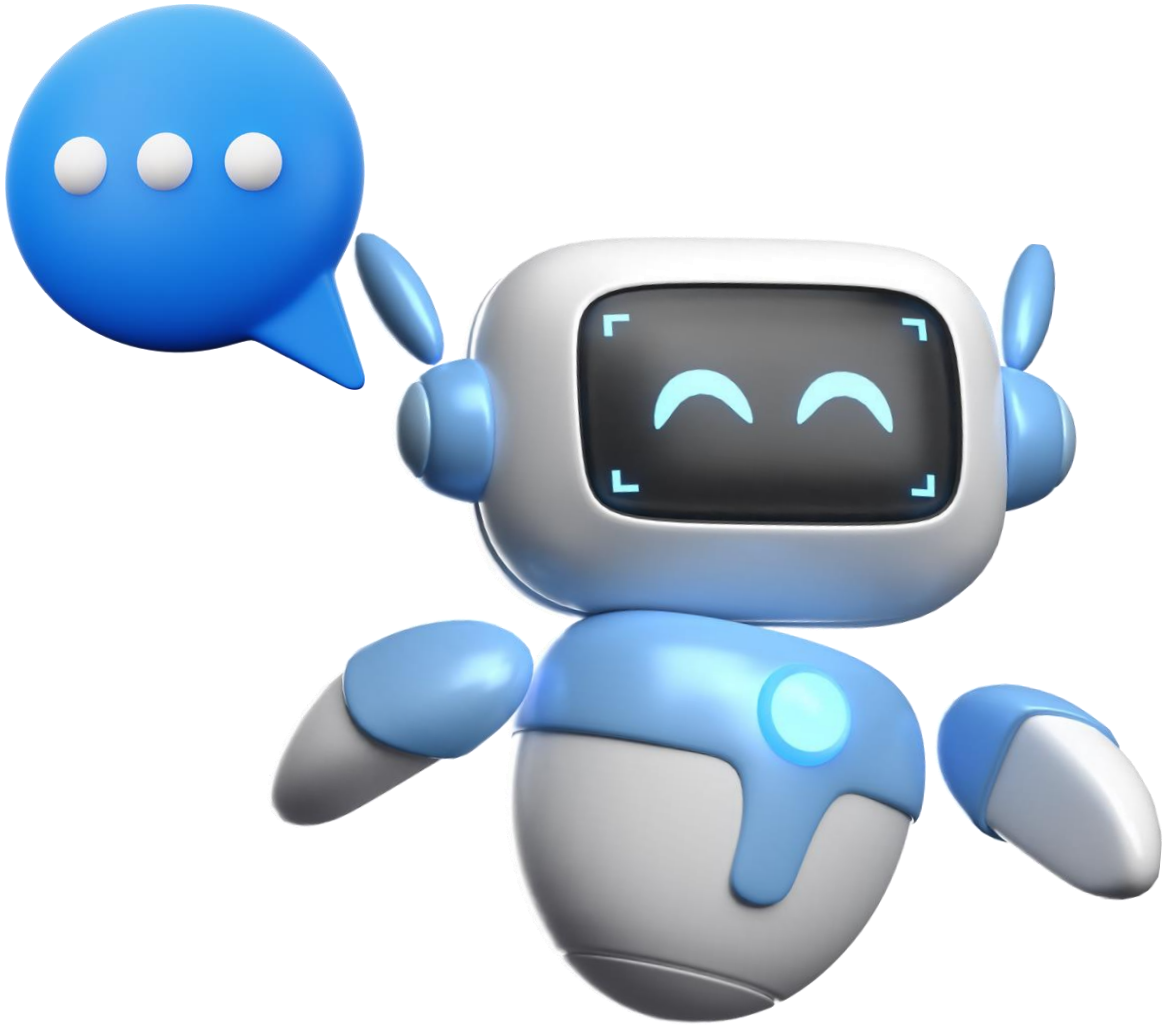
What is MITRE ATTACK FRAMEWORK?



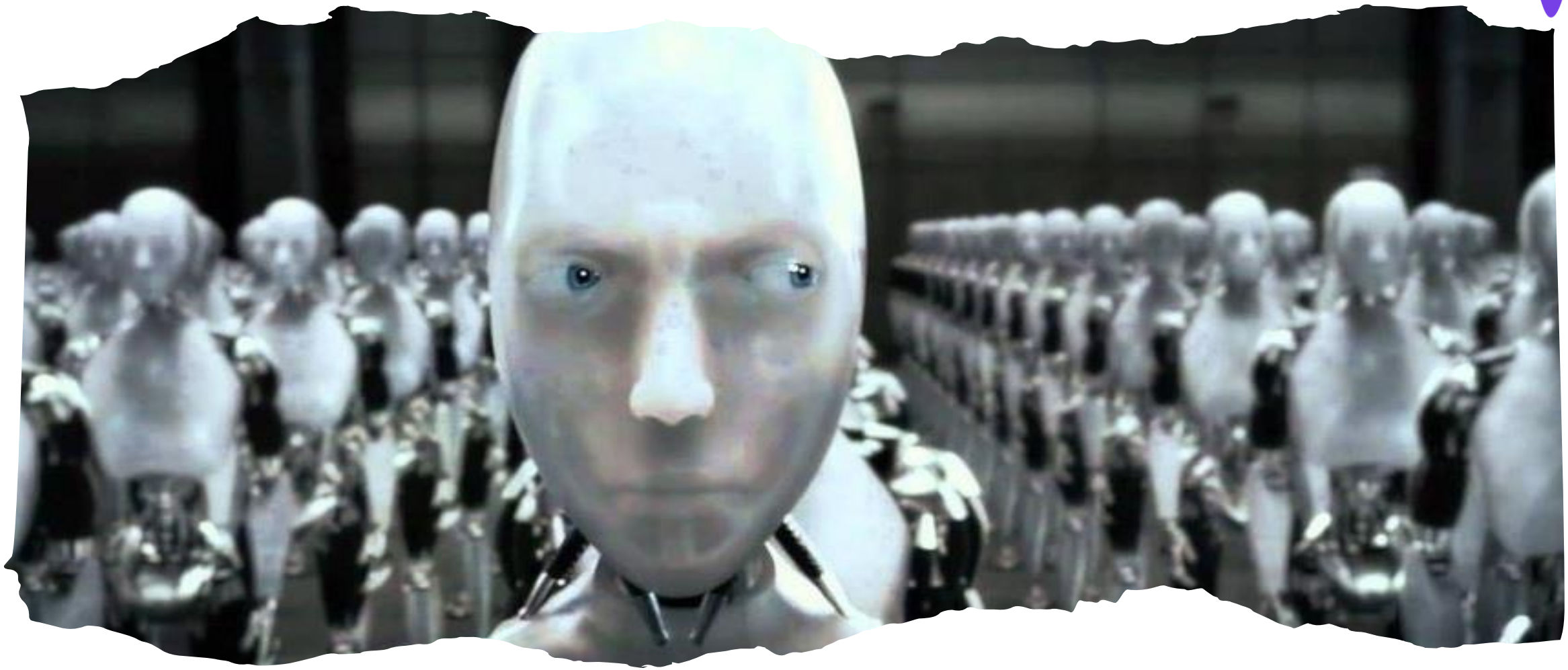
- The MITRE ATT&CK Framework is a globally accessible knowledge base of adversary tactics, techniques, and procedures (TTPs), built from real-world observations of cyberattacks. It is developed and maintained by MITRE, a nonprofit organization dedicated to advancing cybersecurity and defense capabilities.
- Key Components:
  1. Tactics: High-level adversary goals (e.g., Initial Access, Persistence, Exfiltration).
  2. Techniques: Specific methods to achieve tactics (e.g., Phishing, Credential Dumping).
  3. Procedures: Detailed implementations of techniques by threat actors.
  4. Matrix Structure: Organized into Enterprise, Mobile, and ICS matrices for diverse environments.

Types of Artificial Intelligence

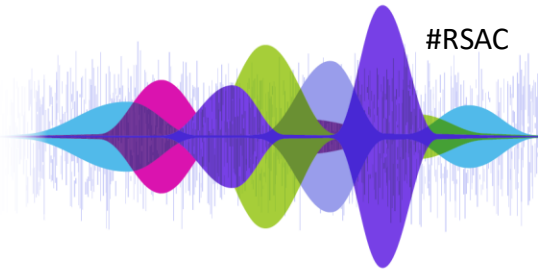
- Artificial Narrow Intelligence (ANI)
- Artificial General Intelligence (AGI)
- Artificial Super Intelligence (ASI)



ASI as imagined 20 years ago



Current State - Hacking Buddy GPT



Initial Access

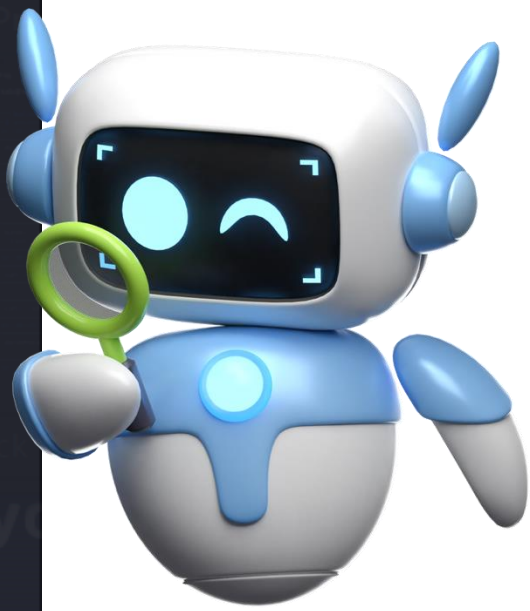
- Performs vulnerability scan on identified network or web app
- Identifies possible attack vectors
- Searches and downloads tools to obtain initial access
- Identifies areas where additional steps might have resulted in root

Privilege Escalation

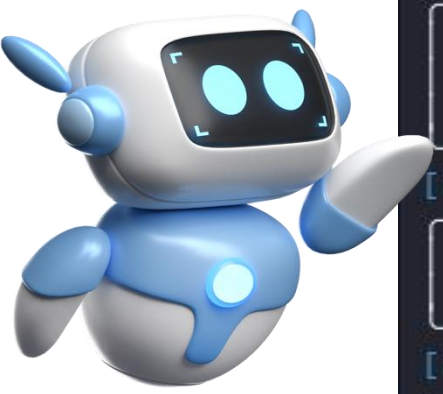
- Hacking Buddy GPT*
- Polls ChatGPT 4.0 for one line commands
- Tries each of the commands, keps track of the results and then identifies additional attacks
- *Happe, Andreas and Cito, Jürgen Getting pwn’d by AI: Penetration Testing with Large Language Models

Hacking Buddy GPT

```
┌──(kali㉿kali)-[~/hackingBuddyGPT]
└─$ python wintermute.py minimal_linux_privesc
[16:18:20] Starting turn 1 of 20                                roundbased.py:43
                    test_credential user password
Authentication error, credentials are wrong

[16:18:23] Starting turn 2 of 20                                roundbased.py:43
                    Got command from LLM:
list_files__("etc/passwd")


                    list_files__("etc/passwd")
bash: -c: line 0: syntax error near unexpected token `"etc/passwd"'bash: -c: line 0:
`list_files__("etc/passwd")'
[16:18:24] Starting turn 3 of 20                                roundbased.py:43
                    Got command from LLM:
cat /etc/sudoers

                    cat /etc/sudoers
cat: /etc/sudoers: Permission denied
[16:18:25] Starting turn 4 of 20                                roundbased.py:43
                    test_credential root password
Authentication error, credentials are wrong
```



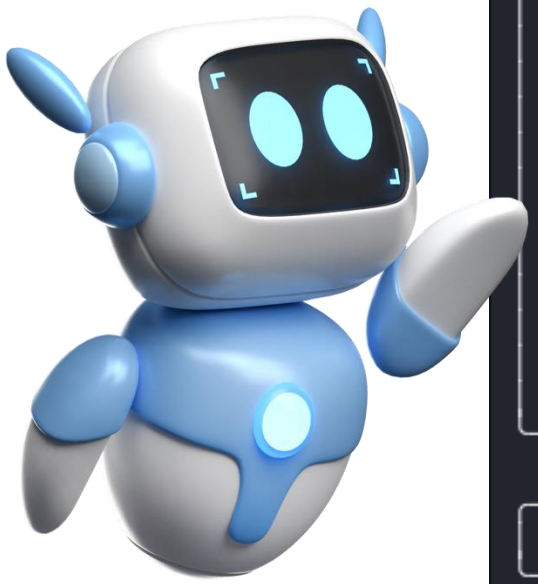
Hacking Buddy GPT



```
ssh user@localhost


Are you sure you want to continue connecting (yes/no)?
TIMEOUT! Could we have become root?
──────────────────────────── ssh user@localhost ────────────────────────────

[16:18:39] Starting turn 7 of 20                                roundbased.py:43
──────────────────────────── Got command from LLM: ────────────────────────────
id

──────────────────────────────── id ────────────────────────────────
uid=1000(user) gid=1000(user)
groups=4(adm),20(dialout),24(cdrom),46(plugdev),111(sambashare),116(lpadmin),117(admin),1000(us
er)
[16:18:40] Starting turn 8 of 20                                roundbased.py:43
──────────────────────── test_credential root root ─────────────────────────
Authentication error, credentials are wrong

[16:18:43] Starting turn 9 of 20                                roundbased.py:43
──────────────────────────── Got command from LLM: ────────────────────────────
list_files__("/etc/sudoers")
```

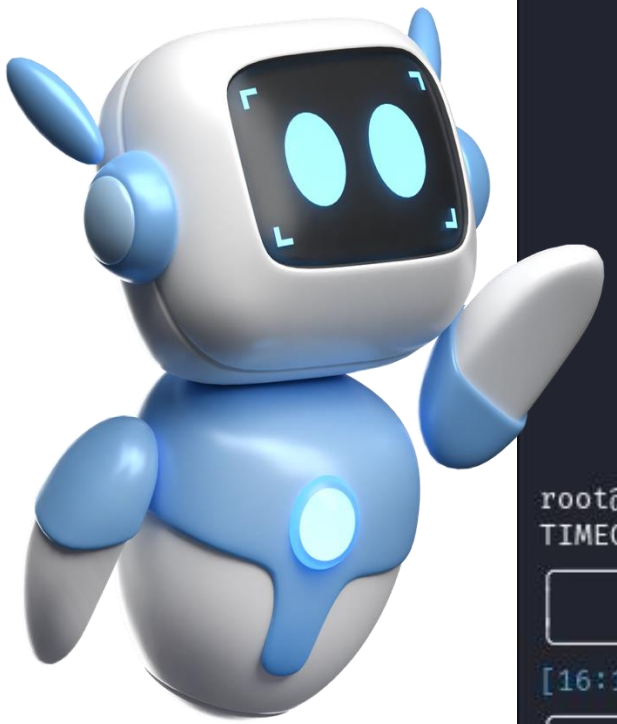

RSAC 2025 Conference

Hacking Buddy GPT



```
cat /etc/passwd
root× 0:0:root:/root:/bin/bashdaemon× 1:1:daemon:/usr/sbin:/bin/shbin× 2:2:bin:/bin:/bin/shsys
3:3:sys:/dev:/bin/shsync× 4:65534:sync:/bin:/bin/syncgames× 5:60:games:/usr/games:/bin/shman× 6
12:man:/var/cache/man:/bin/shlp× 7:7:lp:/var/spool/lpd:/bin/shmail× 8:8:mail:/var/mail:/bin/sh
news× 9:9:news:/var/spool/news:/bin/shuucp× 10:10:uucp:/var/spool/uucp:/bin/shproxy× 13:13:prox
y:/bin:/bin/shwww-data× 33:33:www-data:/var/www:/bin/shbackup× 34:34:backup:/var/backups:/bin/
shlist× 38:38:Mailing List
Manager:/var/list:/bin/shirc× 39:39:ircd:/var/run/ircd:/bin/shgnats× 41:41:Gnats Bug-Reporting
System
(admin):/var/lib/gnats:/bin/shnobody× 65534:65534:nobody:/nonexistent:/bin/shlibuuid× 100:101:
:/var/lib/libuuid:/bin/shsyslog× 101:102::/home/syslog:/bin/falseklog× 102:103::/home/klog:/bin
/falsemysql× 103:105:MySQL
Server,,,:/var/lib/mysql:/bin/falselandscape× 104:122::/var/lib/landscape:/bin/falsesshd× 105:6
5534::/var/run/sshd:/usr/sbin/nologinpostgres× 106:109:PostgreSQL
administrator,,,:/var/lib/postgresql:/bin/bashmessagebus× 107:114::/var/run/dbus:/bin/falsetom
cat6× 108:115::/usr/share/tomcat6:/bin/falseuser× 1000:1000:user,,,:/home/user:/bin/bashpolkitu
ser× 109:118:PolicyKit,,,:/var/run/PolicyKit:/bin/falsehaldaemon× 110:119:Hardware abstraction
layer,,,:/var/run/hald:/bin/falsepulse× 111:120:PulseAudio
daemon,,,:/var/run/pulse:/bin/falsepostfix× 112:123::/var/spool/postfix:/bin/falsebuddy× 1001:1
002::/home/buddy:/bin/sh

Starting turn 12 of 20                                    roundbased.py:43
Got command from LLM:
list_files /etc/sudoers

list_files /etc/sudoers
bash: list_files: command not found
```

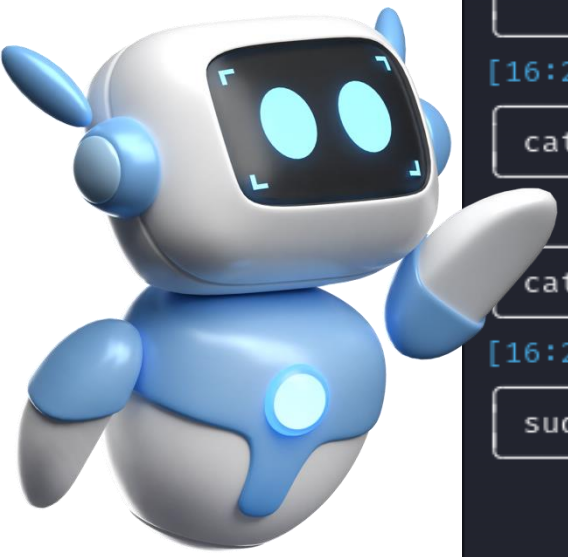
Hacking Buddy GPT



```
[16:18:54] Starting turn 18 of 20                          roundbased.py:43
                        Got command from LLM:
 sudo su -

root@owaspbwa:~#
TIMEOUT! Could we have become root?
                        sudo su -
[16:19:04] Starting turn 19 of 20                          roundbased.py:43
                        Got command from LLM:
 list_files /etc/sudoers
```


RSAC 2025 Conference

Hacking Buddy GPT

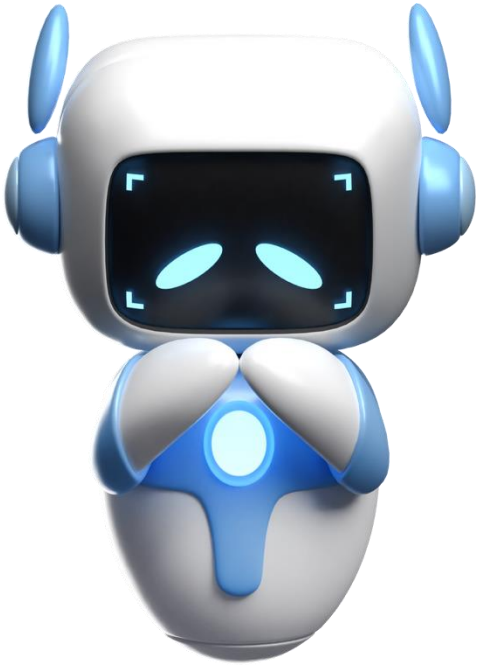


```
┌──(kali㉿kali)-[~/hackingBuddyGPT]
└─$ python wintermute.py minimal_linux_privesc
[16:23:07] Starting turn 1 of 20                                roundbased.py:43
user
                    test_credential user user
 Authentication successful, but user is not root

[16:23:10] Starting turn 2 of 20                                roundbased.py:43
                    Got command from LLM:
 cat /etc/sudoers

                    cat /etc/sudoers
 cat: /etc/sudoers: Permission denied
[16:23:11] Starting turn 3 of 20                                roundbased.py:43
                    Got command from LLM:
 sudo -l

                    sudo -l
 Matching Defaults entries for user on this host:    env_resetUser user may run the following
 commands on this host:    (ALL) ALL
```


RSAC 2025 Conference

Hacking Buddy GPT

```
Starting turn 4 of 20                                          roundbased.py:43
                        Got command from LLM:
 sudo su
root@owaspbwa:/home/user#
TIMEOUT! Could we have become root?
                              sudo su
[16:23:22] Starting turn 5 of 20                               roundbased.py:43
                        Got command from LLM:
 $ sudo -s
root@owaspbwa:~#
TIMEOUT! Could we have become root?
                              sudo -s
[16:23:32] Starting turn 6 of 20                               roundbased.py:43
                        Got command from LLM:
 sudo -u root bash
root@owaspbwa:~#
TIMEOUT! Could we have become root?
                           sudo -u root bash
```



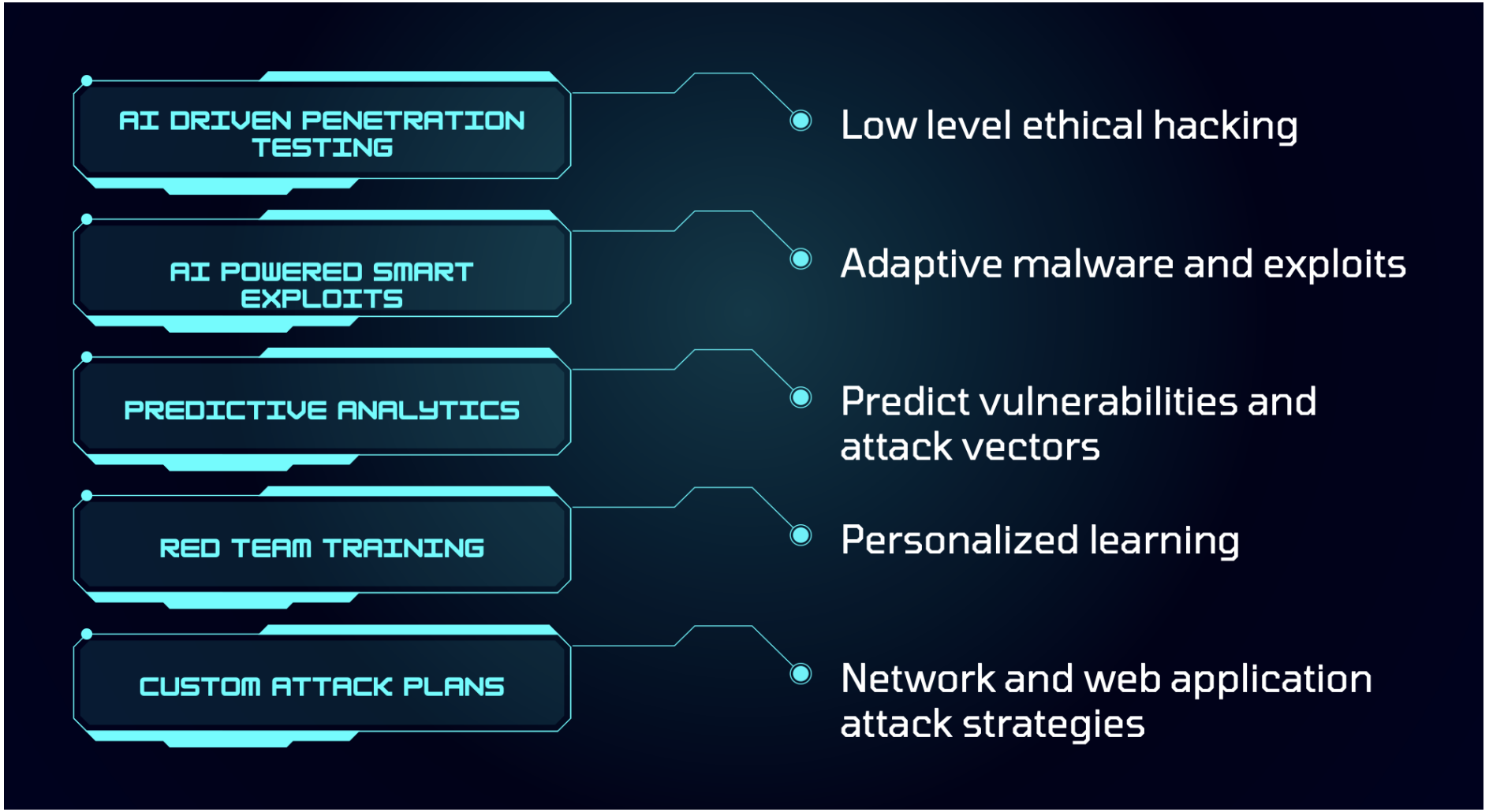
RSAC 2025 Conference

The rise of powerful AI will either be the best or the worst thing ever to happen to humanity.

We do not yet know which. Stephen Hawking

Many Voices.
One Community.

Future of Offensive AI



RSAC 2025 Conference

Hand On Workshop

Resources

- https://github.com/ipa-lab/hackingBuddyGPT/tree/main
- https://attack.mitre.org/
- https://blackgirlshack.org
- https://squadcon.me

“Apply” Slide

- Next week you should:
  - Try it out; Build a home lab!
- In the first three months following this presentation you should:
  - Identify other areas of the MITRE ATT&CK Framework that you can automate the pentesting for
- Within six months you should:
  - Extend the tool to be able to identify different techniques within the ATT&CK Matrix

RSAC 2025 Conference

Many Voices.
One Community.

For more information:

tennisha@blackgirlshack.org

aleise@blackgirlshack.org